matches the definition of the structure "known city", which is not shown in the figures),

plus

A malus of −10 associated with the structure #7 "street address".

In the case of pattern 2, we obtain for A a value of −10, the value of the malus associated with structure #7, since the elements of the pattern "4 Apple Pies" do not match any of the definition items with a bonus.

In operation **114**, A is then compared to a predetermined threshold, here 0. Accordingly, pattern 1 is confirmed since A=5>0 (operation **116**), whereas pattern 2 is rejected since A=−10<0 (operation **118**).

Hence, with the inventive weighting scheme, contrary to the prior art, false positives such as "4 Apple Pies" are spotted and discarded. The inventive method therefore renders pattern searching more effective and accurate.

The invention claimed is:

1. A machine-implemented method for identifying patterns in text using structures defining types of patterns which are to be identified, wherein a structure comprises one or more definition items, the method comprising:

assigning a fixed weighting to each structure and each definition item in each structure, each fixed weighting being an integer multiple of the same integer;

searching the text for a pattern to be identified on the basis of a particular structure, a pattern being provisionally identified if it matches the definition given by said particular structure;

in a provisionally identified pattern, determining definition items making up said particular structure that have been identified in the provisionally identified pattern;

combining the fixed weightings of the determined definition items and the fixed weighting of the particular structure to a single quantity;

assessing whether the single quantity fulfils a given condition; and

depending on the result of said assessment, rejecting or confirming the provisionally identified pattern.

2. The method of claim **1**, wherein the given condition corresponds to the single quantity being above or below a given threshold.

3. The method of claim **1**, wherein the single quantity is obtained by combining the fixed weightings using one or more arithmetic operations.

4. The method of claim **3**, wherein the arithmetic operation is a summation over all fixed weightings, the single quantity being the sum of all the fixed weightings.

5. The method of claim **1**, each fixed weighting taking the form of either a bonus in the form of a positive integer, or a malus in the form of a negative integer.

6. The method of claim **5**, wherein a structure or definition item is assigned a bonus if it is well-defined, and a malus if it is ambiguous.

7. A program storage medium having a program stored therein for causing a data processing system to execute the method of any one of claims **1-4** or **5-6**.

8. An apparatus for identifying patterns in text using structures defining types of patterns which are to be identified, wherein a structure comprises one or more definition items, the apparatus comprising:

a processing system coupled to memory;

means for assigning a fixed weighting to each structure and each definition item in each

structure, each fixed weighting being an integer multiple of the same integer; means for searching the text for a pattern to be identified on the basis of a particular

structure, a pattern being provisionally identified if it matches the definition given by said particular structure;

in a provisionally identified pattern, means for determining definition items making up said particular structure that have been identified in the provisionally identified pattern;

means for combining the fixed weightings of the determined definition items and the

fixed weighting of the particular structure to a single quantity;

means for assessing whether the single quantity fulfils a given condition; and depending on the result of said assessment, means for rejecting or confirming the

provisionally identified pattern.

9. The apparatus of claim **8**, wherein the given condition corresponds to the single quantity being above or below a given threshold.

10. The apparatus of claim **8**, wherein the single quantity is obtained by the means for combining the fixed weightings using one or more arithmetic operations.

11. The apparatus of claim **8**, wherein the arithmetic operation is a summation over all fixed weightings, the single quantity being the sum of all the weightings.

12. The apparatus of claim **8**, each fixed weighting taking the form of either a bonus in the form of a positive integer, or a malus in the form of a negative integer.

* * * * *